

АНАЛИЗ РЕЗУЛЬТАТОВ ТЕСТИРОВАНИЯ И КАЧЕСТВА ТЕСТОВ ПРИ СЕРТИФИКАЦИИ СПЕЦИАЛИСТОВ ПО НЕРАЗРУШАЮЩЕМУ КОНТРОЛЮ

© 2024 г. В.В. Муравьев^{1,2,*}, Л.В. Волкова¹, О.В. Муравьева^{1,2}, С.А. Мурашов¹

¹ФГБОУ ВО ИжГТУ имени М.Т. Калашникова, Россия 426069 Ижевск, ул. Студенческая, 7

²ФГБУН «УдмФИЦ УрО РАН», Россия 426067 Ижевск, ул. им. Татьяны Барамзиной, 34

E-mail: *vmuraviev@mail.ru

Поступила в редакцию 29.01.2024; после доработки 12.02.2024

Принята к публикации 13.02.2024

Анализируются результаты квалификационных экзаменов при сертификации персонала по неразрушающему контролю акустическими, магнитным и вихретоковым методами объектов железнодорожного транспорта. Проанализировано влияние уровня образования, стажа работы, возраста кандидатов на результат прохождения квалификационных экзаменов. Проведен анализ результатов общего и специального экзамена по неразрушающему контролю вихретоковым методом за 2021 и 2022 годы на выявление экстремальных вопросов теста. По методу Раша проведен анализ соответствия трудности тестовых заданий уровню подготовленности кандидатов на сертификацию. Построены характеристические кривые трудности тестовых заданий и характеристические кривые обученности тестируемых. По результатам анализа выявлены тестовые задания с завышенным и заниженным логитами трудности задания по сравнению с логитом обученности кандидатов. Оценено соответствие трудности теста уровню подготовленности персонала и равномерность распределения заданий по трудности. Рассчитан требуемый логит обученности кандидата на сертификацию для успешного прохождения тестов общего и специального экзамена.

Ключевые слова: сертификация персонала, неразрушающий контроль, метод Раша, латентный параметр, квалификационный экзамен.

ANALYSIS OF TEST RESULTS AND TEST QUALITY DURING CERTIFICATION OF NON-DESTRUCTIVE TESTING SPECIALISTS

V.V. Murav'ev^{1,2,*}, L.V. Volkova¹, O.V. Murav'eva^{1,2}, S.A. Murashov¹

¹Kalashnikov Izhevsk State Technical University, Russia 426069, Izhevsk, Studencheskaya, 7

²Udmurt Federal Research Center Ural Branch Russian Academy of Sciences, Russia 426067 Izhevsk, st. them.

Tatyana Baramzina, 34

E-mail: *vmuraviev@mail.ru

This article analyzes the results of personal certification for eddy current testing of railway facilities. The influences of education level, the work experience, the candidate age on the qualifying examinations results are considered. The conformity between test task difficulties and certification candidate training is analyzed with Rasch method. The characteristic curves of test task difficulties and the characteristic curves of candidate training levels are constructed. The results general and special examinations of 2020 and 2021 are analyzed to identify extreme test questions. The research results enable to assess the candidate training and the test task difficulties through latent parameters measured as logit unit. The analysis identified test tasks with overstated and understated task difficulty logits compared with the logits of the candidate training, assessed the correspondence between test difficulties and training candidate level, and assessed a distribution uniformity of the tasks difficulty. The required training logit of candidates for the certification is calculated for successful passing general and special exam tests. The getting results give a basis for further analysis of the tasks complexity, identification the reasons for the insufficiency of the auditory material familiarization for the tasks with low rate of response, for making adjustments to the consulting sessions, as well as the content and form of the test.

Keywords: Non-destructive inspection, eddy current testing, staff certification, qualification exam, Rasch method.

DOI: 10.31857/S0130308224020052

ВВЕДЕНИЕ

Тестирование — это стандартизированная форма контроля, позволяющая оценить знания, умения и навыки экзаменуемого. Тестирование широко используется при сертификации и аттестации различных категорий персонала технических, педагогических и других специальностей.

Сертификация персонала в области неразрушающего контроля (НК) объектов железнодорожного транспорта является процедурой проверки квалификации персонала, проводимой уполномоченной организацией, согласно ГОСТ Р ИСО 9712—2023 «Контроль неразрушающий. Квалификация и сертификация персонала. Основные требования».

Сертификация персонала в области НК проводится в целях подтверждения достаточности теоретической и практической подготовки, опыта, компетентности персонала, т.е. его профессиональных знаний, навыков, мастерства, и предоставления права на выполнение работ по одному или нескольким видам (методам) НК. Оценка знаний персонала при их сертификации на уровень квалификации по видам и методам НК является одной из наиболее трудоемких и ответственных операций. На достоверность и объективность оценки качества знаний персонала может влиять ряд факторов, обусловленных когнитивными и личностными трудностями кандидата на сертификацию при взаимодействии с экзаменаторами, методикой проведения собеседования и др. ГОСТ Р ИСО 9712—2023 регламентирует при сертификации персонала тестовую форму проведения общего и специального экзамена. К безусловным достоинствам компьютерного тестирования при сертификации относятся объективность, независимость, единые требования оценки знаний кандидата на сертификацию, оперативность сертификации, охват тестовыми заданиями всего требуемого программой сертификации объема знаний физических основ и технологий метода НК.

Следует иметь в виду, что на результаты тестирования влияют стаж работы, уровень образования, средний возраст испытуемого [1—3]. Содержание тестовых вопросов может быть неравномерным и неоднородным, при этом способность усвоения различных разделов программы сертификации зависит не только от уровня подготовки специалиста к тестированию и качества преподавания в учебных центрах, но и от проработанности самих тестовых заданий.

В статье представлен анализ результатов тестирования при проведении квалификационных экзаменов в процессе сертификации персонала по вихретоковому контролю (ВТК) с использованием метода Раша [4—7], проводимых в экзаменационном центре по неразрушающему контролю Ижевского государственного технического университета имени М. Т. Калашникова (ЭЦНК).

Цель работы — проанализировать влияние уровня образования, стажа работы, возраста кандидатов на результат квалификационных экзаменов по НК, оценить соответствие трудности тестовых заданий по вихретоковым методам и уровня подготовленности экзаменуемого кандидата к сертификации, дать рекомендации по модернизации материалов тестирования.

ИСХОДНЫЕ ПОЛОЖЕНИЯ И ПОСТАНОВКА ЗАДАЧИ

Анализ проведен на основании архива личных дел ЭЦНК — Экзаменационного центра по неразрушающему контролю ИжГТУ имени М. Т. Калашникова (рег. № РОСС RU.0001.03Н300.ЭЦ-38 от 25 февраля 2021 г.). Исследовалось влияние возраста, стажа работы, уровня образования кандидатов на сертификацию на результаты тестирования. Качество тестовых материалов рассмотрено на примере анализа результатов тестирования по вихретоковому виду НК за 2021—2022 гг.

Согласно ГОСТ Р ИСО 9712—2023, для подтверждения общей компетенции лица, проводящего контроль, орган по сертификации должен провести квалификационный экзамен, на котором оцениваются общие, специальные и практические знания и навыки кандидата. Для прохождения общего и специального экзамена кандидату требуется дать ответы на экзаменационные вопросы с вариантами ответов (тестовые вопросы). Любое лицо, успешно сдавшее квалификационные экзамены, получает заявленный уровень сертификации, соответствующий продемонстрированным компетенциям при проведении НК. (В заявке на сертификацию работодатель указывает первый или второй квалификационный уровень сертификации направляемого претендента.) В рамках своих компетенций персонал первого уровня имеет право подготавливать оборудование для проведения контроля, проводить контроль, записывать и классифицировать результаты контроля, составлять отчеты о результатах контроля.

Персонал второго уровня дополнительно разрабатывает технологические карты для проведения контроля, определяет и проверяет установки оборудования, проводит и руководит контролем, оценивает результаты контроля.

Персонал третьего уровня должен обладать компетенциями, которые позволяют ему принимать полную ответственность за аппаратуру контроля, разрабатывать инструкции и процедуры по НК, определять конкретные методы контроля, методики и инструкции по НК, подлежащие применению, участвовать в подготовке персонала или в качестве экзаменатора при проведении квалификационных экзаменов.

Тестовые материалы составлены на основе утвержденных программ сертификации и включают по общему экзамену следующие основные разделы: общие вопросы НК, материаловедение, физические основы, методы и аппаратура вихретокового контроля.

Таблица 1

Количество анализируемых тестовых вопросов и протоколов тестирования по вихретоковому контролю

Вид экзамена	Количество анализируемых вопросов		Количество анализируемых протоколов	
	2021 г.	2022 г.	2021 г.	2022 г.
Общий	201	217	23	15
Специальный	57	95	24	17

Базы данных ЭЦНК включают в себя как тестовые задания, используемые при проведении общего экзамена, так и результаты тестирования в виде протоколов. Общее количество анализируемых вопросов и протоколов указано в табл. 1.

Тесты общего экзамена по определенному виду контроля содержат по 50 заданий для каждого сертифицируемого с весом каждого задания 2 %. Задания для теста выбираются из базы данных случайным образом, но с разделением на темы и без повторов. Тестирование каждого кандидата на сертификацию продолжается не более 120 мин. По окончании тестирования формируются протоколы тестирования, содержащие результаты ответов на каждый вопрос и соответствующий им процент правильных ответов. Результат теста общего экзамена не менее 70 % (количество правильных ответов 35 и более) засчитывается как успешно сданный. Кандидаты, успешно прошедшие тестирование по общему экзамену, допускаются к собеседованию по 2-3 несданным вопросам, требующим развернутых ответов.

Тест специального экзамена содержит 20 заданий для каждого кандидата с весом каждого задания 5 % и продолжается не более 60 мин. Несдавшим специальный экзамен считается кандидат, получивший по результатам тестирования менее 70 % правильных ответов, т. е. ответивший правильно менее чем на 14 вопросов. Собеседование по специальному экзамену включает ответы на 3 вопроса по технологии контроля объекта согласно заявленному производственному сектору.

В программе и тестах специального экзамена рассматриваются вопросы технологии вихретокового контроля деталей вагонов и средства контроля, нормативные документы на проведение контроля, типы выявляемых дефектов, критерии браковки, этапы проведения контроля, настройка оборудования для контроля, схемы сканирования, оценка качества.

При анализе результатов тестирования с использованием логистических моделей возможно оценить не только количество и долю правильных ответов, но и ряд других (латентных — недоступных для прямого наблюдения) параметров. Латентные параметры определяются косвенно как функции от доступных для прямого наблюдения индикаторных переменных, в качестве которых выступают непосредственно испытуемые и задания тестов (в виде их порядковых номеров или индексов). Значениями индикаторных переменных могут быть правильность или неправильность ответа конкретного испытуемого на конкретное тестовое задание.

Латентные параметры измеряются в безразмерных условных единицах (логитах), легко переводимых в любую интервальную шкалу и в дальнейшем используемых в широком классе процедур статистического анализа [8—10].

При анализе качества тестовых заданий с позиций соответствия их трудности уровню подготовленности кандидатов на сертификацию использовалась математическая модель Раша [10—13], согласно которой анализируется статистика количества правильных и неправильных ответов кандидатов на вопросы теста и рассчитываются логит подготовленности (обученности) i -го тестируемого θ_i и логит трудности j -го вопроса β_j в относительных единицах измерения одной шкалы:

$$\theta_i = \ln(C_i/I_i), \quad (1)$$

$$\beta_j = \ln(W_j/R_j), \quad (2)$$

где C_i и I_i — количество соответственно правильных и неправильных ответов i -го кандидата на вопросы теста; R_j и W_j — количество соответственно правильных и неправильных ответов на j -й вопрос теста.

Латентные параметры обученности и трудности позволяют определить вероятность $P_i(\theta)$ того, что кандидат с уровнем подготовки θ_i справится с заданием со сложностью β :

$$P_i(\theta) = \{1 + \exp[-1,7(\theta_i - \beta)]\}^{-1}, \quad (3)$$

и вероятность $P_j(\beta)$ того, что с заданием сложностью β_j справится тестируемый с уровнем подготовки θ :

$$P(\beta_j) = \{1 + \exp[-1,7(\theta - \beta_j)]\}^{-1}. \quad (4)$$

В частности, если разность между уровнем подготовленности участника тестирования и трудностью тестового задания составляет 1 логит, то вероятность верного выполнения этим испытуемым такого задания равна 73 %.

Знание латентных параметров позволяет также оценить надежность теста, характеризующую точностью и устойчивостью результатов тестирования при многократном применении, что ранее было показано на примерах магнитного и акустико-эмиссионного методов [14, 15].

Анализ связи характеристик сертифицируемых кандидатов и результатов тестирования

Для оценки качества тестовых материалов и процедуры сертификации в ЭЦНК был проведен анализ результатов тестов по нескольким критериям:

- распределение кандидатов, выдержавших испытания на II уровень сертификации по различным видам НК;
- распределение кандидатов по присвоенному уровню сертификации;
- распределение кандидатов, справившихся с общим и специальным экзаменом;
- распределение кандидатов по стажу работы и уровню образования;
- распределение кандидатов по возрасту.

Также произведена оценка логитов трудности заданий и уровня подготовленности специалистов. Анализ результатов тестов произведен по четырем видам контроля, по которым проводится квалификационный экзамен в экзаменационном центре по неразрушающему контролю: ультразвуковому, магнитному, вихретоковому, акустико-эмиссионному.

Результаты деятельности ЭЦНК по сертификации персонала на I, II, III уровни по всем аккредитованным видам НК за период 8 лет показаны на рис. 1.

Общее количество кандидатов по всем видам НК, выдержавших испытания на II уровень сертификации, составляет 1765 человек; кроме того, 124 человека сдали на I уровень и 11 человек — на III уровень сертификации (рис. 1а). Существенное преобладание специалистов, получающих II уровень сертификации, связано с требованиями работодателей, которым необходим персонал с допуском на проведение НК и выдачу заключений о его результатах, но без обязанности по составлению технологических инструкций НК.

На рис. 1б представлено распределение кандидатов, выдержавших квалификационные испытания на II уровень сертификации по различным видам НК: видно, что по вихретоковому контролю

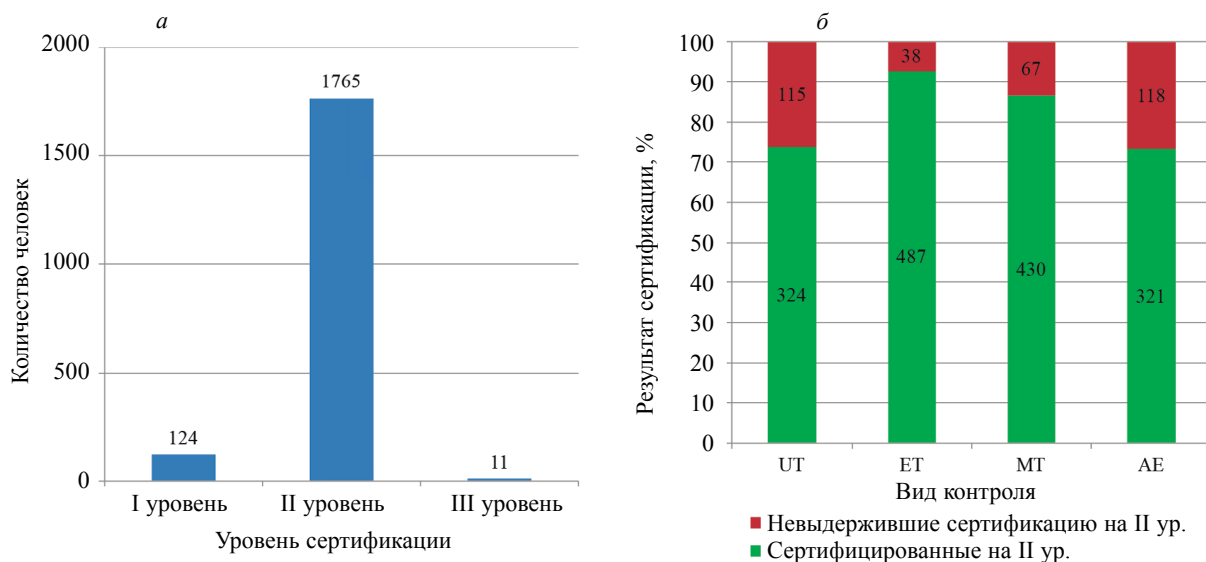


Рис. 1. Результаты присвоения уровней сертификации (а) и II уровня сертификация по различным видам неразрушающего контроля (б).

(ЕТ) специалисты получают II уровень в 93 % случаев, по магнитному контролю (МТ) — в 87 % случаев, по ультразвуковому (УТ) и акустико-эмиссионному контролю (АЕ) — по 73 % случаев.

Результаты тестирования в зависимости от производственного стажа кандидата

На результат тестирования по общему и специальному экзамену может влиять стаж работы кандидата на сертификацию, который для получения II уровня сертификации по ВТК должен составлять не менее 12 месяцев. Для оценки данного фактора построено распределение результатов общего и специального экзамена в зависимости от стажа работы (рис. 2а). Получено, что с общим экзаменом, независимо от стажа работы, кандидаты справляются с вероятностью успеха около 85 %. Со специальным экзаменом на знание технологии контроля кандидаты с низким стажем (1 год) справляются с вероятностью 78 %, а кандидаты со стажем более 15 лет справляются с вероятностью около 90 %.

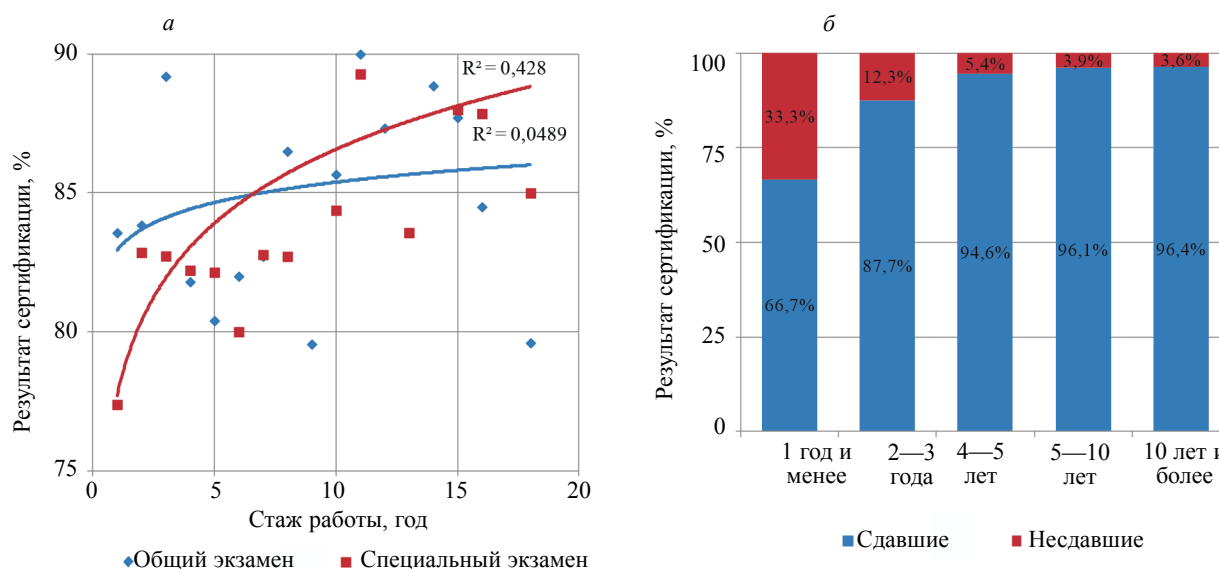


Рис. 2. Результат прохождения общего и специального экзаменов в зависимости от стажа работы (а) и сертификации специалистов в зависимости от стажа работы (б).

Общий процент сдачи аттестационных экзаменов для кандидатов со стажем 1 год составляет 66,7 %, со стажем 2-3 года — 87 %, а для кандидатов со стажем работы от 4 лет и более — стремится к 100 % (рис. 2б).

Распределение сертифицируемых кандидатов по образованию

Распределение кандидатов, приезжающих для прохождения квалификационных экзаменов, по имеющемуся у них образованию показано на рис. 3: среднее профессиональное и среднее специальное образование имеют 62 % кандидатов, высшее образование — 27 %, среднее — 9 %, незаконченное высшее образование — 2 %.

Оценка результатов присвоения кандидатам уровня сертификации в зависимости от образования показывает, что кандидаты с высшим образованием получают уровень сертификации в 90 % случаев, со средним специальным — в 88 %, а со средним образованием — в 85 % (рис. 4). То есть, уровень образования в меньшей степени влияет на получение 2 уровня сертификации по сравнению со стажем работы, так как в процессе работы кандидаты на сертификацию проходят периодическое повышение квалификации и техническую учебу, что повышает их базовый уровень знаний, умений и навыков.

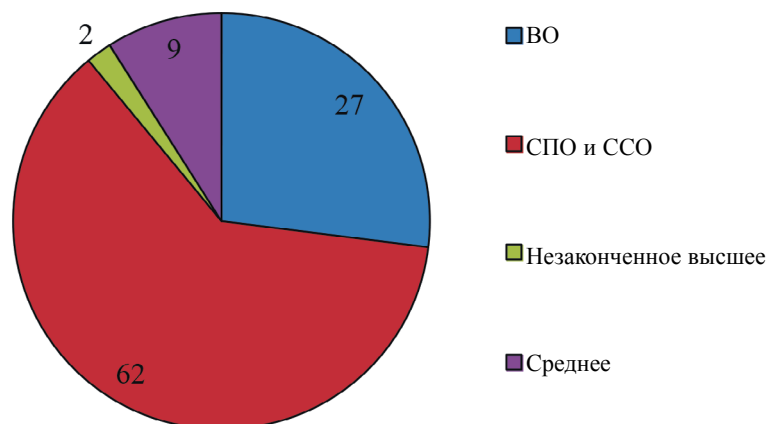


Рис. 3. Распределение кандидатов на сертификацию по уровню образования.

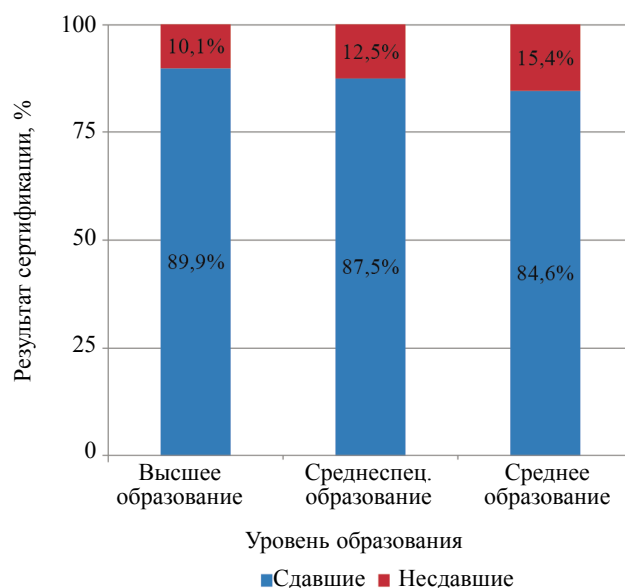


Рис. 4. Результат сдачи квалификационных экзаменов в зависимости от уровня образования.

Распределение сертифицируемых кандидатов по возрасту

Средний процент сдачи квалификационных экзаменов повышается с возрастом кандидатов (рис. 5): для кандидатов 1971 г. рождения и старше он составляет 95 %, для кандидатов 1981—1990 гг. рождения — 85 %, что также является хорошим результатом, для кандидатов 1991—1998 гг. рождения — 74 %, что всего лишь на 4 % выше порогового значения, при котором экзамен считается сданным. Данные показатели также обусловлены стажем работы кандидатов и накопленными знаниями, умениями и навыками в процессе производственной деятельности, а также предыдущих процедур сертификаций по НК.

На результат прохождения квалификационных экзаменов в большей степени влияет качество тестового материала как измерительного инструмента, то есть совместимость вопросов теста при измерении уровня знаний кандидатов.

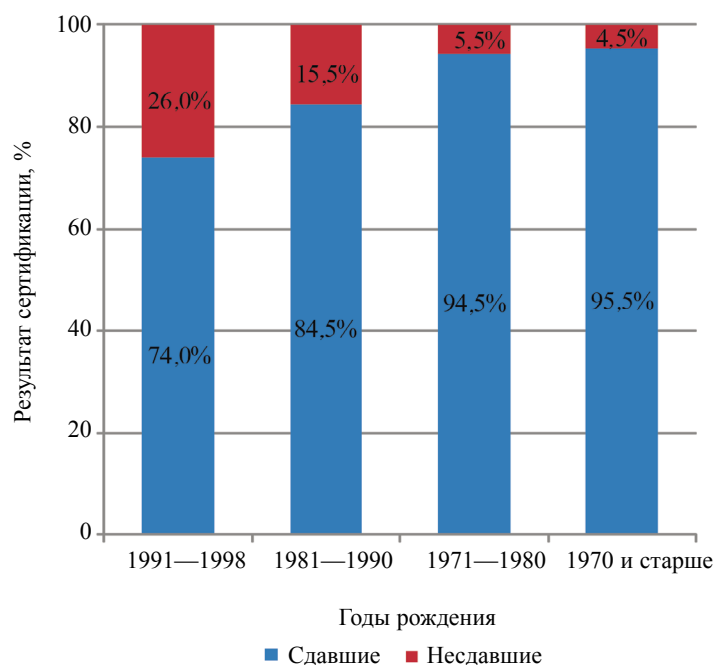


Рис. 5. Результат прохождения квалификационных экзаменов в зависимости от возраста испытуемых.

Анализ результатов общего экзамена

На рис. 6 представлены гистограммы результатов тестирования группы испытуемых по общему экзамену вихретокового контроля, проведенного в 2021 (а) и в 2022 гг. (б). Данные ранжированы по возрастанию результата кандидата с условным номером протокола (первый условный номер протокола присвоен специалисту с минимальным результатом тестирования). Согласно рис. 6а, в 2021 г. все 23 кандидата были допущены, из них 12 кандидатов (52 %) имели результат тестирования не ниже 90 %. Согласно рис. 6б, в 2022 г. из 15 тестируемых кандидатов 5 человек (30 %) имели результат тестирования не ниже 90 % и 1 кандидат не справился с тестом.

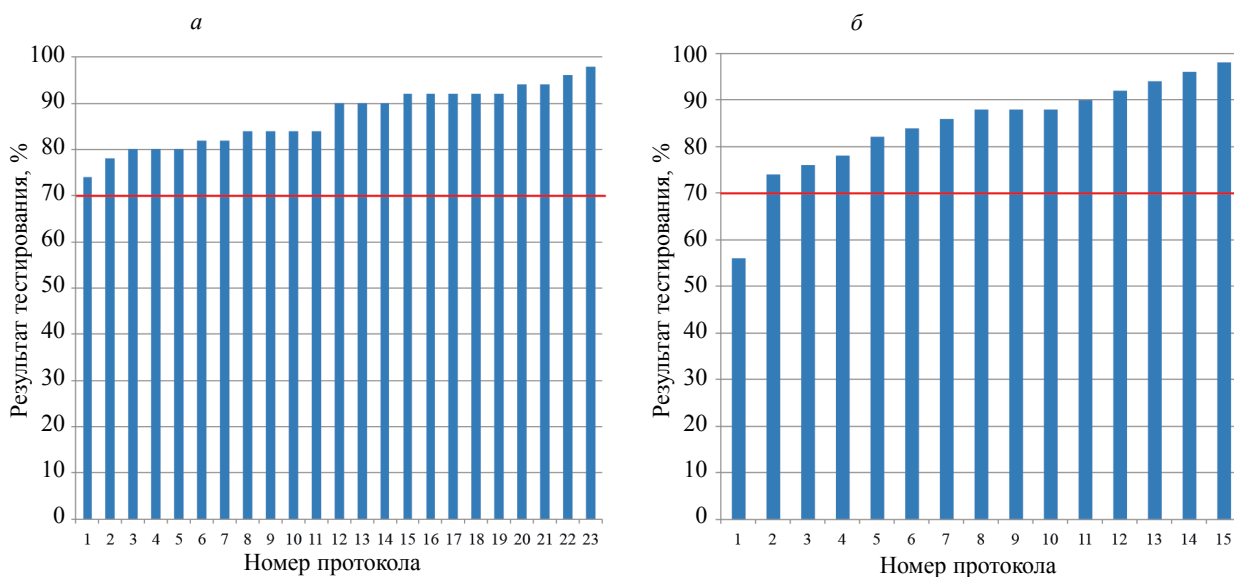


Рис. 6. Результаты тестирования по общему экзамену вихретокового контроля в 2021 (а) и 2022 (б) гг.

Следует отметить, что в 2022 г. тестовые задания были пересмотрены с целью более равномерного охвата разделов программы, улучшения формулировок, усиления сложности вопросов. В связи с группировкой вопросов по темам, распределение вероятностей попадания вопросов в тест является неравномерным и варьируется от 5 до 60 %. В рассматриваемом тесте за 2021 г. из банка вопросов общего теста хотя бы один раз был задан 201 вопрос, в тесте за 2022 г. были использованы 217 вопросов.

Распределение результатов ответов по отдельным вопросам тестов по общему экзамену вихревого контроля иллюстрирует рис. 7. Гистограммы ранжированы по возрастанию процента правильных ответов на вопрос с условным номером (первый условный номер протокола соответствует вопросу, на который был дан минимальный процент правильных ответов). Видно, что в тесте 2021 г. 70 % испытуемых дали верные ответы на 163 задания из 201 (81 % заданий), а в тесте за 2022 г. 70 % испытуемых дали верные ответы на 165 заданий из 217 (76 % заданий). В тестах выявлены экстремально по сложности задания, на которые не было получено верных ответов или, наоборот, все ответы были верными. Если исключить из рассмотрения задания, заданные всего один раз, то в тесте 2021 г. выявлено 2 задания без верных ответов (сложные) и 106 заданий со стопроцентным положительным результатом (простые), в тесте 2022 г. — 4 и 109 заданий соответственно. При этом количество полученных ответов варьируется от 2 до 3 для экстремально сложных вопросов и от 2 до 13 для экстремально простых. Номера экстремально сложных вопросов в тестах 2021 и 2022 гг. не совпадают. Количество экстремально простых вопросов приближается к 50 %, что требует их пересмотра в сторону усложнения либо исключения из теста.

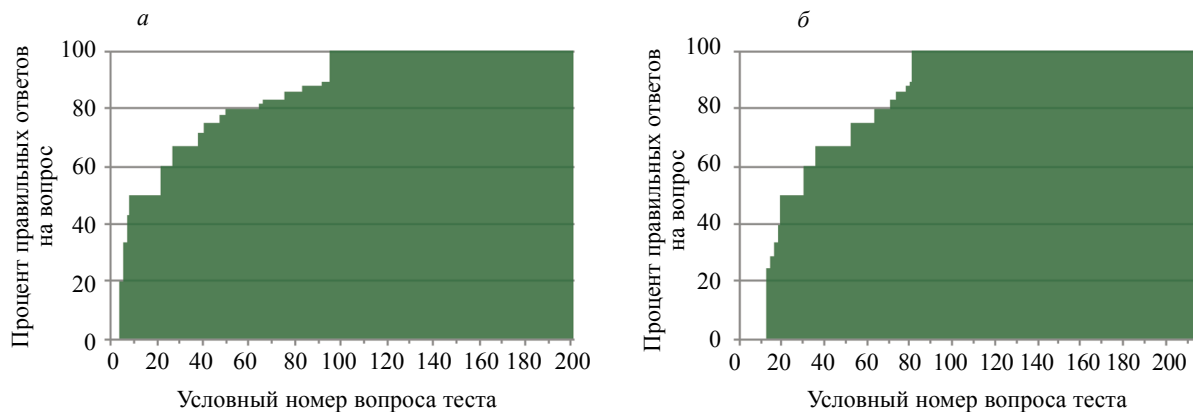


Рис. 7. Распределение правильных ответов по отдельным вопросам теста общего экзамена в 2021 (а) и 2022 (б) гг.

Гистограммы распределения значений латентных параметров уровня подготовленности кандидатов и трудности заданий в логитах представлены на рис. 8. Отрицательным значениям логитов трудности заданий и обученности кандидатов соответствуют наиболее легкие задания и низкий уровень подготовки кандидатов. Напротив, положительным значениям логитов — наиболее сложные задания и хорошо подготовленные кандидаты. Известно, что средняя трудность заданий не должна отличаться от среднего уровня подготовленности более чем на 0,5 логита [6, 7]. В противном случае оценка уровня знаний кандидатов будет производиться с различной точностью в зависимости от диапазона трудностей заданий.

Из анализа результатов следует, что средний уровень подготовленности кандидатов (среднее значение логита обученности) в 2022 г. составил $\theta = 0,96$ логита, что на 0,2 логита меньше среднего уровня трудности задания (среднего значения логита трудности задания $\beta = 1,16$), тогда как в 2021 г. средний уровень логита подготовленности кандидатов составил $\theta = 0,95$, что на 0,34 логита больше среднего уровня трудности заданий $\beta = 0,64$. Последнее подтверждает, что уровень подготовленности кандидатов соответствует уровню трудности заданий. Диапазон изменения уровня подготовленности кандидатов находился в диапазоне $\theta = -0,5—2,5$ логита в 2022 г. и $\theta = 0—2,5$ логита в 2021 г. При этом уровень трудности заданий изменялся в диапазоне $\beta = -1,0—3$ логита и полностью покрывал интервал изменения уровня подготовленности кандидатов как в 2021 г., так и в 2022 г.

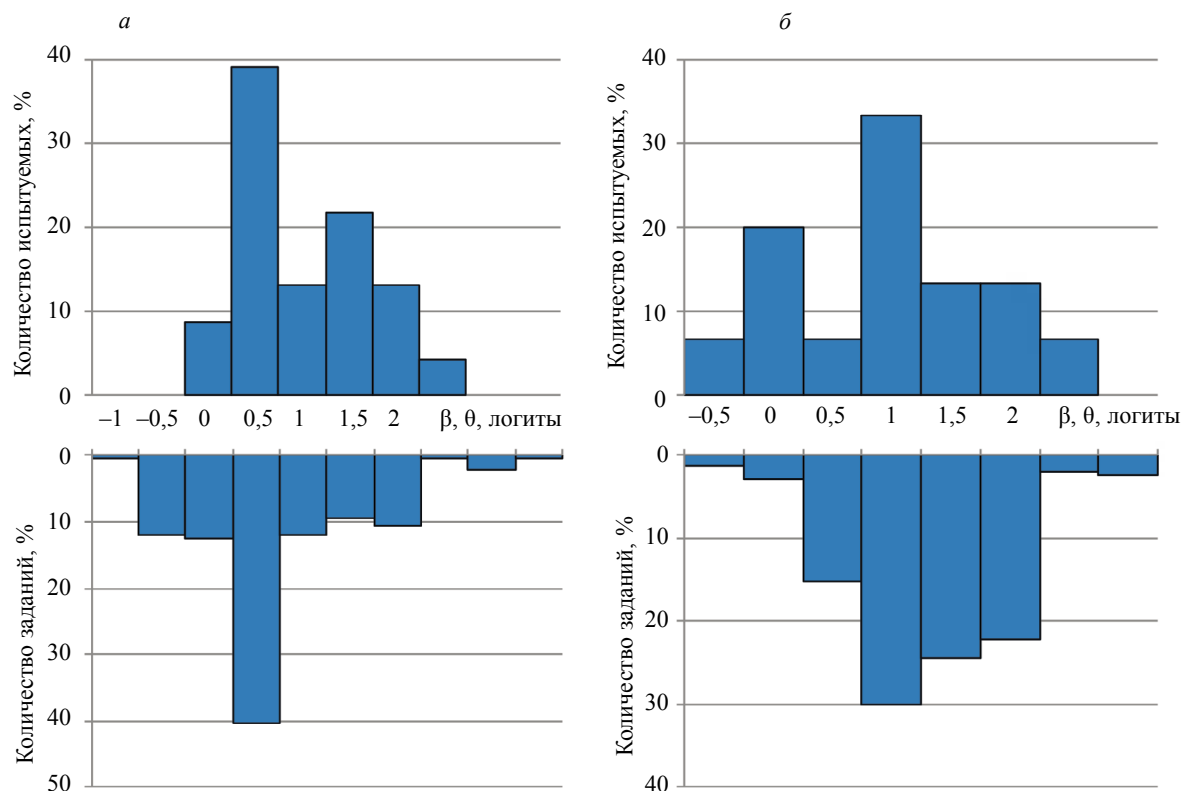


Рис. 8. Результаты распределения уровня подготовленности кандидатов и трудности заданий общего экзамена в 2021 (а) и в 2022 (б) гг.

Для достоверной оценки уровня подготовленности кандидатов при анализе качества тестовых заданий необходимо, чтобы гистограмма распределения уровня трудностей заданий была близка к равномерному закону распределения и охватывала весь диапазон. В исследуемом тесте распределение трудностей заданий достаточно равномерно. Смещение максимума распределения на 0,5—1 логит вправо вполне оправданно, так как процент сдачи должен составлять 70 %. Анализируя гистограмму распределения трудностей заданий, можно определить, какие задания необходимо пересмотреть с точки зрения увеличения или уменьшения их трудности. В исследуемом тесте к наиболее легким тестовым заданиям относятся задания с логитом трудности $\beta = -1,0$ — $0,5$, к заданиям средней трудности — с логитом $\beta = 1,0$ — $2,0$, к наиболее трудным — с логитом $\beta > 2,5$.

Примеры экстремальных вопросов теста (с минимальными и максимальными логитами трудности), выделенные в 2021 и 2022 гг., представлены в табл. 2.

Такого рода экстремальные задания требуют пересмотра формулировки либо исключения из теста. Для обеспечения равномерного распределения трудностей заданий необходимо уменьшить количество заданий с логитом трудности от 1,0 и выше и увеличивать количество заданий с логитом трудности 0,5.

Анализ результатов специального экзамена

База данных, используемая при проведении специального экзамена, включает тестовые задания в количестве 142 вопроса. Сдавшим тестирование при специальном экзамене считается кандидат, правильно ответивший не менее чем на 14 вопросов (70 %) из предложенных 20. Кандидаты, успешно прошедшие тестирование, допускаются на собеседование, где экзаменационная комиссия задает 2 или 3 вопроса, касающиеся методики работы и требующие развернутого ответа. Анализ результатов тестирования по специальному экзамену 24 кандидатов за 2021 г. и 12 кандидатов за 2022 г., претендовавших на II уровень сертификации по вихретоковому контролю, представлен на рис. 9.

Таблица 2

Экстремальные вопросы общего экзамена по ВТК

Логит	Номер вопроса	Формулировка вопроса	Варианты ответа
Тест 2021 г.			
-0,74	31	К источнику синусоидального напряжения частотой ω подключены последовательно соединенные конденсатор емкостью C и катушка индуктивностью L . В каком случае ток в цепи будет максимальным?	1. $\omega L = \omega C$ 2. $\omega L = 1/(\omega C)$ 3. $\omega L \gg \omega C$ 4. $\omega L \ll \omega C$
3,58	145	Чем обусловлено изменение индуктивного сопротивления вихретокового преобразователя?	1. Потерями энергии на вихревые токи за счет нагрева объекта контроля вихревыми токами 2. Изменением потокоцепления индуктивной катушки за счет действия вихревых токов 3. Изменением емкости катушки
Тест 2022 г.			
-0,15	25	Активное сопротивление обмотки определяется:	1. Материалом провода обмотки 2. Длиной провода 3. Поперечным сечением провода 4. Всеми указанными факторами
+2,99	132	К чему приводит уменьшение плотности вихревых токов в поверхностном слое?	1. Уменьшается влияние шероховатости и уменьшается чувствительность к мелким поверхностным дефектам 2. Увеличивается влияние шероховатости и уменьшается чувствительность к мелким поверхностным дефектам 3. Увеличивается влияние шероховатости и увеличивается чувствительность к мелким поверхностным дефектам 4. Уменьшается влияние шероховатости и увеличивается чувствительность к мелким поверхностным дефектам

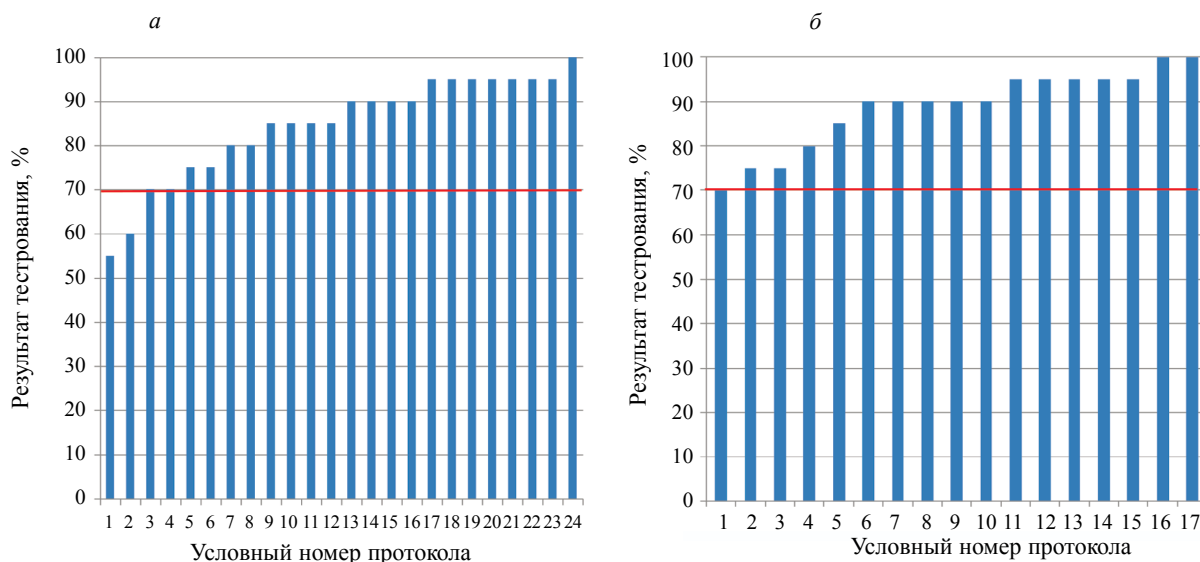


Рис. 9. Результаты тестирования кандидатов на сертификацию по специальному экзамену вихретокового контроля в 2021 (а) и 2022 (б) гг.

Согласно рис. 9, среди 24 кандидатов, претендующих на II уровень квалификации по вихретоковому контролю в 2021 г., 1 кандидат прошел тест с результатом 100 %, 2 кандидата набрали менее

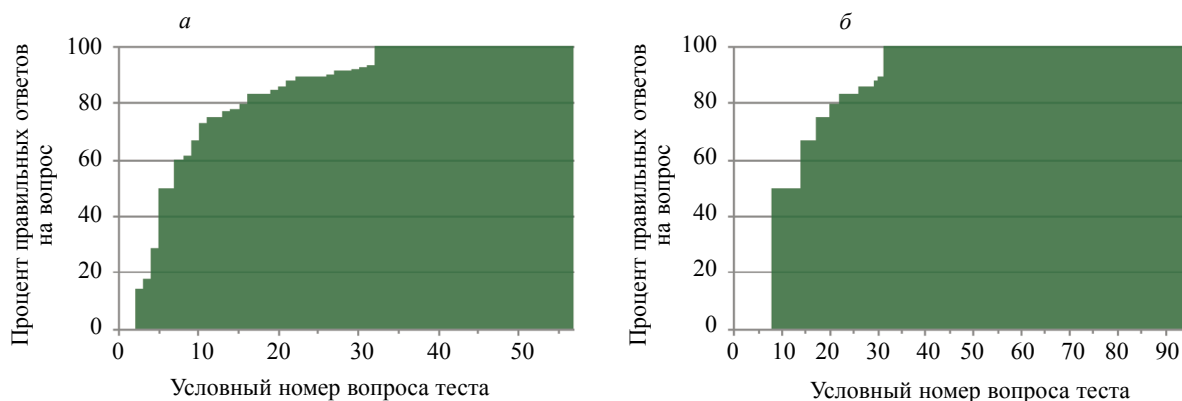


Рис. 10. Распределение правильных ответов по отдельным вопросам теста специального экзамена по вихретоковому контролю в 2021 (а) и 2022 (б) гг.

70 % и не были допущены к собеседованию. В 2022 г. из 17 тестируемых кандидатов все допущены к собеседованию, из них 2 прошли тест со 100 % результатом. Анализ результатов тестирования по отдельным вопросам теста специального экзамена (рис. 10) показывает, что процент правильных ответов на вопросы находится в тесте 2021 г. в интервале 15—100 %, в тесте 2022 г. — в интервале 50—100 %. Банк вопросов состоит из 146 вопросов. Из них в тесты хотя бы один раз попали 56 вопросов в тесте 2021 года и 94 вопроса в тесте 2022 года. Сто процентов кандидатов на сертификацию ответили на 25 вопросов в тесте 2021 г. и на 64 вопроса в тесте 2022 г. Правильный ответ не был дан ни одним из кандидатов на 2 вопроса в тесте 2021 г., на 8 вопросов в тесте 2022 г., что свидетельствует об увеличении трудности после переработки тестов.

Анализ гистограмм распределения значений латентных параметров уровня подготовленности кандидатов на сертификацию и трудности заданий (рис. 11) показывает, что средний уровень подготовленности кандидатов в 2021 г. на 0,29 логита больше уровня трудности заданий, а в 2022 г. — на 0,16 логита меньше.

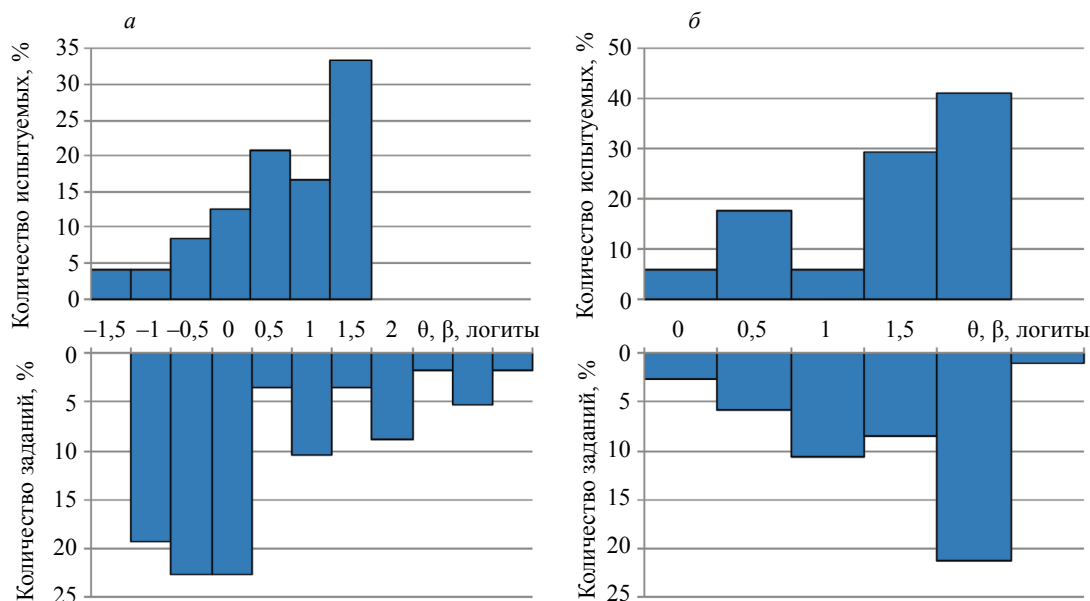


Рис. 11. Распределение уровня подготовленности кандидатов на сертификацию и трудности заданий специального экзамена по вихретоковому контролю в 2021 (а) и 2022 (б) гг.

В 2021 г. уровень трудности заданий изменялся в диапазоне 1,0—3,5 логита, в 2022 г. диапазон изменения трудности заданий составлял 0—2,5 логита, при этом равномерность в распределении

трудностей заданий отсутствовала. Уровень трудности заданий не покрывает уровень подготовленности кандидатов в 2021 г., что говорит о трудности теста. В 2022 г. уровень трудности полностью покрывал интервал изменения уровня подготовленности кандидатов, что позволяет говорить о соизмеримости уровня трудности заданий и уровня подготовленности кандидатов на сертификацию.

Примеры экстремальных вопросов, выявленных в тестах специального экзамена по вихретоковому контролю, представлены в табл. 3.

Таблица 3

Экстремальные вопросы специального экзамена по вихретоковому контролю

Логит	Номер вопроса	Формулировка вопроса	Варианты ответа
Тест 2021 г.			
-0,97	11	В каких дефектоскопах используются роторные вихретоковые преобразователи?	1) ВД-12НФМ, ВД-12НФ, ВД-15НФ 2) ВД-12НФМ 3) ВД-113.5, ВД-113 4) ВД-12НФ
4,01	32	С помощью какого стандартного образца устанавливают порог чувствительности при контроле валика тягового хомута?	1) СОП—НО-038, используя ИД-2 2) СОП—НО-038, используя ИД-1 3) СОП—НО-038, используя ИД-5 4) СОП—НО-038, используя ИД-3
Тест 2022 г.			
0	6	Что относится к помехам вихретокового вида неразрушающего контроля?	1) Шероховатость поверхности детали 2) Локальные изменения размера зерна металла объекта контроля 3) Локальные изменения электромагнитных свойств металла 4) Локальные изменения твердости объекта контроля 5) Изменение зазора между ВП и металлической поверхностью детали 6) Изменение кривизны контролируемой поверхности 7) Верны 1), 3), 5), 6) 8) Верны 1), 2), 3), 6)
2,87	51	Какова шероховатость рабочей поверхности отраслевого стандартного образца ОСО-Г-043:	1) Rz 160 2) Rz 20 3) Rz 320 4) Ra 20

Общее качество тестовых материалов, используемых при сертификации кандидатов на сертификацию по вихретоковому контролю, можно оценить по характеристическим кривым трудности заданий и обученности кандидатов.

Анализ характеристической кривой теста для общего экзамена (рис. 12а) показывает, что для успешной сдачи экзамена кандидаты на сертификацию должны были иметь логит обученности не менее 1,5 в 2021 г. и не менее 2 в 2022 г. Средняя сложность тестовых заданий (рис. 12б) в 2021 г. составила 0,9 логита, вероятность ответа на тест составила 0,55, а в 2022 г. — 1,2 логита и вероятностью ответа на тест 0,49. Высокий процент сдачи по общему экзамену связан с большим количеством легких вопросов, на которые все кандидаты дают правильные ответы.

Из характеристических кривых тестовых заданий общего экзамена (рис. 13а) видно, что задания неравномерно покрывают требуемый диапазон уровней подготовленности испытуемых (θ от -5 до +7), особенно выпадают из равномерного распределения задания с трудностью +3,59 логит. Сравнение характеристических кривых уровня подготовленности (рис. 13б), показывает что логит обученности специалистов лежит в достаточно широком диапазоне от 1,13 до 3,89, однако все кривые равномерно покрывают уровень трудности заданий от -4 до +7 логита трудности заданий.

Анализ характеристической кривой теста специального экзамена (рис. 14а) показывает, что для успешной сдачи теста логит обученности должен быть не менее 1,5 в 2021 г. и не менее 1,8 в 2022 г. Средняя сложность тестовых заданий в 2021 и в 2022 гг. составила 0,9 логита, а вероятность ответа на тест в 2021 г. составила 0,35, в 2022 г. — 0,55 (рис. 14б).

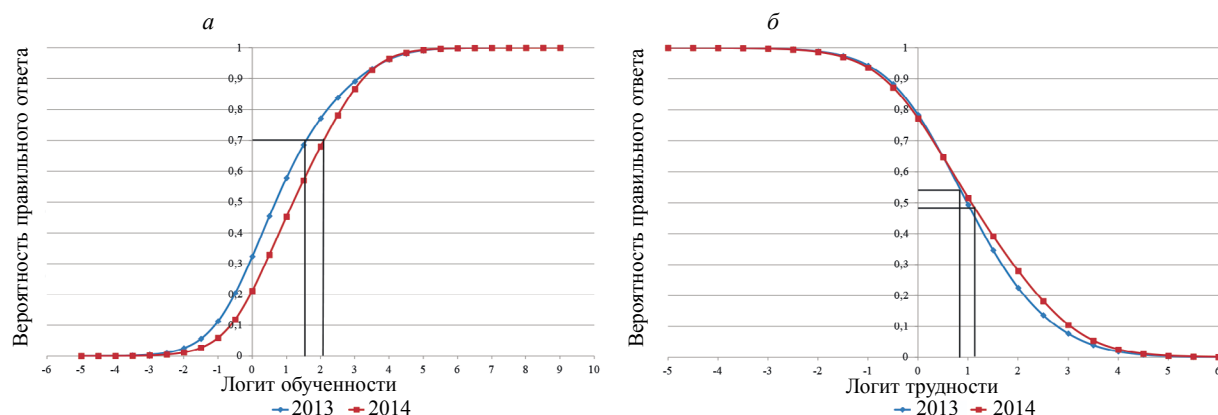


Рис. 12. Характеристические кривые теста общего экзамена (а) и подготовленности кандидатов на сертификацию (б) в 2021—2022 гг.

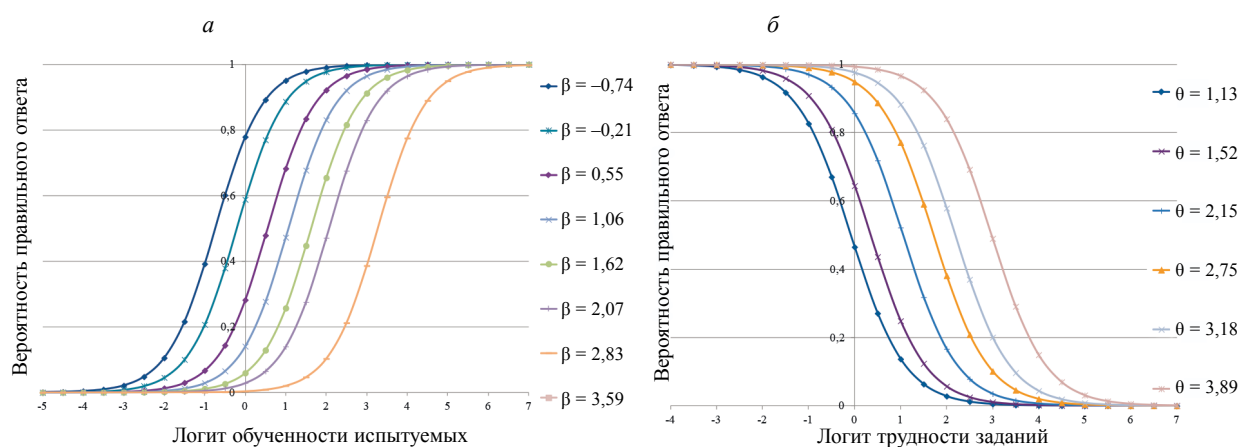


Рис. 13. Характеристические кривые тестовых заданий общего экзамена (а) и испытуемых (б).

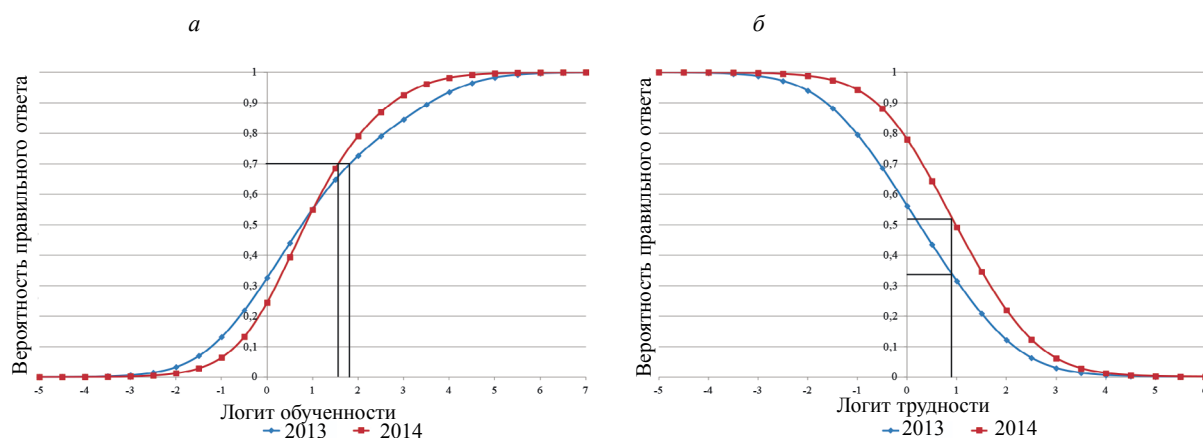


Рис. 14. Характеристические кривые теста специального экзамена (а) и испытуемых (б).

ВЫВОДЫ

Анализ результатов прохождения сертификационных испытаний на различные виды неразрушающего контроля показал, что большинство специалистов имеют высокий уровень подготовки к

экзаменам, связанный с уровнем образования, приобретенными навыками и опытом работы в НК не менее одного года.

Для исключения предвзятости при тестировании в экзаменационном центре кандидаты проходят собеседование, в процессе которого оценивается уверенность кандидата на сертификацию в применении основных понятий, знании технологий контроля, а также применении знаний в решении практических задач контроля.

Анализ качества заданий согласно модели Раша позволяет:

- выявить и исключить тестовые задания с экстремальными значениями;
- оценить соответствие трудности теста уровню подготовленности тестируемых специалистов;
- определить диапазон варьирования трудности тестовых заданий;
- оценить равномерность распределения заданий по трудности.

Анализ результатов тестовых материалов при сертификации специалистов по вихретоковому виду НК показал:

- наличие отклонений в гистограммах распределения уровня трудности тестовых заданий от равномерного закона, особенно для тестов специального экзамена, что требует увеличения числа заданий с низкими и высокими логитами трудности, где наблюдаются минимумы функции;
- необходимость усиления подготовки специалистов по вопросам, соответствующим минимальным логитам обученности.

Полученные результаты дают основания для дальнейшего анализа заданий по сложности, причин недостаточного усвоения материала в вопросах с низким процентом ответов, возможных внесений корректировки в консультационные занятия, а также в содержание и форму теста.

СПИСОК ЛИТЕРАТУРЫ

1. *Berrio Angela I., Gomez-Benito J., Arias E.M.* Developments and Trends in Research on Methods of Detecting Differential Item Functioning // *Educational Research Review*. 2020. V. 31. P. 100340. DOI:10.1016/j.edurev.2020.100340
2. *Boon H.J., Lewthwaite B.* Development of an Instrument to Measure a Facet of Quality Teaching: Culturally Responsive Pedagogy // *International Journal of Educational Research*. 2015. V. 72. P. 38—58. DOI: 10.1016/j.ijer.2015.05.002
3. *Aghekyan R.* Validation of the SIEVEA Instrument Using the Rasch Analysis // *International Journal of Educational Research*. 2020. V. 103. P. 101619. DOI:10.1016/j.ijer.2020.101619
4. *Kruit P., Oostdam R., van den Berg E., Schuitema J.* Performance Assessment as a Diagnostic Tool for Science Teachers // *Research in Science Education*. 2020. V. 50. No. 3. P. 1093—117. DOI:10.1007/s11165-018-9724-9
5. *Vaughan B.* A Rasch Analysis of the Revised Study Process Questionnaire in an Australian Osteopathy Student Cohort // *Studies in Educational Evaluation*. 2018. V. 56. P. 144—53. DOI:10.1016/j.stueduc.2017.12.003
6. *Ehrich J.F., Woodcock S., West C.* The Effect of Gender on Teaching Dispositions: A Rasch Measurement Approach // *International Journal of Educational Research*. 2020. V. 99. P. 101510. DOI:10.1016/j.ijer.2019.101510
7. *Zhao Yue, Huen J.M., Chan Y.W.* Measuring Longitudinal Gains in Student Learning: A Comparison of Rasch Scoring and Summative Scoring Approaches // *Research in Higher Education*. 2017. V. 58. No. 6. P. 605—16. DOI:10.1007/s11162-016-9441-z
8. *Sabetova T.V.* The use of latent variable method in human resource management // *Russian Journal of Agricultural and Socio-Economic Sciences*. 2017. V. 66. No. 6. P. 123—29. DOI:10.18551/rjoas.2017-06.14
9. *Van Zile-Tamsen C.* Using Rasch Analysis to Inform Rating Scale Development // *Research in Higher Education*. 2017. V. 58. No. 8. P. 922—33. DOI:10.1007/s11162-017-9448-0
10. *Jääskä E., Aaltonen K.* Teachers' Experiences of Using Game-Based Learning Methods in Project Management Higher Education // *Project Leadership and Society*. 2022. V. 3. P. 100041. DOI:10.1016/j.plas.2022.100041
11. *Chang W.-Ch., Ludlow L., Grudnoff L., Ell F., Haigh M., Hill M.* Measuring the Complexity of Teaching Practice for Equity: Development of a Scenario-Format Scale // *Teaching and Teacher Education*. 2019. V. 82. P. 69—85. DOI: 10.1016/j.tate.2019.03.004
12. *Красова Н.А., Алейникова Н.Е.* Математическое моделирование задачи определения набора контрольных мероприятий с помощью нечеткого метода анализа иерархий и метода, основанного на измерении латентных переменных // *Моделирование, Оптимизация и Информационные Технологии*. 2019. Т. 7. № 3 (26). С. 33. DOI: 10.26102/2310-6018/2019.26.3.024
13. *Корытаева Ю.В., Красова Н.Е., Пешков В.В., Рябов С.В.* Методика организации и оценки качества научно-исследовательских работ курсантов военно-научного общества на основе математической

модели Раша // Вестник Воронежского Государственного Технического Университета. 2020. Т. 16. № 2. С. 65—70. DOI: 10.25987/VSTU.2020.16.2.010

14. *Муравьев В.В., Волкова Л.В., Платунов А.В., Булатова Е.Г., Байкалова Т.Н.* Анализ результатов тестирования специалистов по магнитному виду неразрушающего контроля при сертификации // Контроль. Диагностика. 2015. № 10. С. 36—42. EDN UJZFEL. DOI: 10.14489/td.2015.010.pp.036-042

15. *Муравьев В.В., Муравьева О.В., Волкова Л.В., Богдан О.П., Платунов А.В., Байкалова Т.Н.* Анализ качества тестов при сертификации специалистов по акустико-эмиссионному методу контроля объектов железнодорожного транспорта // В мире неразрушающего контроля. 2015. Т. 18. № 3. С. 66—71. EDN UDJSQX.

REFERENCES

1. *Berrío Ángela I., Gomez-Benito J., Arias E. M.* Developments and Trends in Research on Methods of Detecting Differential Item Functioning // Educational Research Review. 2020. V. 31. P. 100340. DOI:10.1016/j.edurev.2020.100340

2. *Boon H.J., Lewthwaite B.* Development of an Instrument to Measure a Facet of Quality Teaching: Culturally Responsive Pedagogy // International Journal of Educational Research. 2015. V. 72. P. 38—58. DOI:10.1016/j.ijer.2015.05.002

3. *Aghekyan R.* Validation of the SIEVEA Instrument Using the Rasch Analysis // International Journal of Educational Research. 2020. V. 103. P. 101619. DOI:10.1016/j.ijer.2020.101619

4. *Kruit P., Oostdam R., van den Berg E., Schuitema J.* Performance Assessment as a Diagnostic Tool for Science Teachers // Research in Science Education. 2020. V. 50. No. 3. P. 1093—117. DOI:10.1007/s11165-018-9724-9

5. *Vaughan B.* A Rasch Analysis of the Revised Study Process Questionnaire in an Australian Osteopathy Student Cohort // Studies in Educational Evaluation. 2018. V. 56. P. 144—53. DOI:10.1016/j.stueduc.2017.12.003

6. *Ehrich J.F., Woodcock S., West C.* The Effect of Gender on Teaching Dispositions: A Rasch Measurement Approach // International Journal of Educational Research. 2020. V. 99. P. 101510. DOI:10.1016/j.ijer.2019.101510

7. *Zhao Yue, Huen J.M., Chan Y.W.* Measuring Longitudinal Gains in Student Learning: A Comparison of Rasch Scoring and Summative Scoring Approaches // Research in Higher Education. 2017. V. 58. No. 6. P. 605—16. DOI:10.1007/s11162-016-9441-z

8. *Sabetova T.V.* The use of latent variable method in human resource management // Russian Journal of Agricultural and Socio-Economic Sciences. 2017. V. 66. No. 6. P. 123—29. DOI:10.18551/rjoas.2017-06.14

9. *Van Zile-Tamsen C.* Using Rasch Analysis to Inform Rating Scale Development // Research in Higher Education. 2017. V. 58. No. 8. P. 922—33. DOI:10.1007/s11162-017-9448-0

10. *Jääskä E., Aaltonen K.* Teachers' Experiences of Using Game-Based Learning Methods in Project Management Higher Education // Project Leadership and Society. 2022. V. 3. P. 100041. DOI:10.1016/j.plas.2022.100041

11. *Chang W.-Ch., Ludlow L., Grudnoff L., Ell F., Haigh M., Hill M.* Measuring the Complexity of Teaching Practice for Equity: Development of a Scenario-Format Scale // Teaching and Teacher Education. 2019. V. 82. P. 69—85. DOI:10.1016/j.tate.2019.03.004

12. *Krasova N.A., Aleynikova N.E.* Mathematical modeling of the problem of determining a set of control measures using the fuzzy method of analyzing hierarchies and a method based on measuring latent variables // Modeling, Optimization and Information Technologies. 2019. V. 7. No. 3 (26). P. 33. DOI:10.26102/2310-6018/2019.26.3.024

13. *Korypaeva Yu.V., Krasova N.E., Peshkov V.V., Ryabov S.V.* Methodology for organizing and assessing the quality of research work of cadets of the military scientific society based on the Rush mathematical model // Bulletin of Voronezh State Technical University. 2020. V. 16. No. 2. P. 65—70. DOI:10.25987/VSTU.2020.16.2.010

14. *Murav'ev V.V., Volkova L.V., Platunov A.V., Bulatova E.G., Bajkalova T.N.* Analysis of test results mean specialists in magnetic non-destructive testing for the certification // Kontrol. Diagnostika. 2015. No. 10. P. 36—42. EDN UJZFEL. DOI: 10.14489/td.2015.010.pp.036-042.

15. *Muravjev V.V., Muravjeva O.V., Volkova L.V., Bogdan O.P., Platunov A.V., Bajkalova T.N.* Test quality analysis during specialist certification on acoustic emission nondestructive testing of railway objects // In the world of non-destructive testing. 2015. V. 18. No. 3. P. 66—71. EDN UDJSQX.